DOCUMENT RESUME

ED 136 261 CS 203 231

AUTHOR Haack, Dennis G.
TITLE Teaching about Survey Doublespeak.
PUB DATE 76
NOTE 38p.; Paper presented at the Annual Meeting of the National Council of Teachers of English (66th, Chicago, November 25-27, 1976)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Communication Problems; *English Instruction; Higher Education; *Statistical Analysis; *Statistics; *Surveys
IDENTIFIERS *Doublespeak

ABSTRACT

Doublespeak--inflated, involved, and often deliberately ambiguous language--can occur in the interpretation of statistics when a "credible" statistic is misused, when an unsupported number is passed off as a credible statistic, or when the limitations of a statistic are not known. This paper, intended for English teachers who are concerned with both verbal and nonverbal doublespeak, details the ways in which statistical doublespeak can occur in surveys, explains how surveys may be evaluated without getting into the strictly statistical aspects of their bases, and provides examples to aid students in a search for survey doublespeak. Special reference is made to terminology, poor preelection polls, and surveys and surveying procedures. (KS)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED136261

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Teaching about Survey Doublespeak

by

Dennis G. Haack*

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Dennis G. Haack
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

1. INTRODUCTION.

Doublespeak is "inflated, involved, and often deliberately ambiguous language." But what can be more "inflated" or more "involved" or more "ambiguous" than the language of statistics?

This form of nonverbal doublespeak is prevalent in our society because citizens of this country have a general lack of understanding, indeed, a real "fear," of the numbers we call statistics. Because people do not know how to critically evaluate statistics, doublespeak will result when

i) a "credible" statistic is misused,

ii) a totally worthless number is passed off as a credible statistic,

or iii) the limitations of a statistic are not known.

Examples of statistical doublespeak abound in the media: public officials and advertisers furnish a "lions" share. Consider, for example

*The author, a member of the NCTE Committee on Public Doublespeak, is an assistant professor of statistics with the Department of Statistics and consultant with the Tobacco and Health Research Institute at the University of Kentucky.

CS 203 231

i) A statistic may be credible but is misused:

> "Ford administration strategists relied on grossly distorted statistics to scare enough House Republicans into switching sides on the strip mining bill so that the President's veto would be upheld...
>
> "A West Virginia professor told Congressmen that the Administration misused his research in concluding that 36,000 jobs might be lost. In fact, he said, the strip mining bill would increase coal industry employment because more deep mines would be opened." (From a newspaper editorial, mid 1975.)

ii) Doublespeak could also result from the use of a meaningless number. Such was the case when a nationally syndicated columnist claimed in a Memorial Day article that 500,000 Vietnam Vets had attempted suicide. In a later column entitled, "Maybe it's good for the soul," this columnist confesses:

> "...This statistic was derived from an article in Penthouse magazine. The author of that article obtained it from a pamphlet of Twice Born Men, a veterans group in San Francisco, now defunct. That organization's former director, Jack McCloskey, says he got the figure from the National Council of Churches of the U.S.A. The Council disavowed any knowledge of the statistic, which must therefore be considered unsupported. It's publication in this space is regretted."

or iii) The use of a statistic could result in doublespeak if one does not know the limitations of that statistic. As an example let us consider unemployment statistics. For January and February of 1975, a recessionary period, unemployment was 8.2% for each month. The unemployment rate is the percentage

of unemployed people in the "labor force." A person is in the labor force only if he or she has made an attempt to find a job. From January 1975 to February 1975 580,000 individuals dropped out of the labor force, perhaps because they became discouraged and quit looking for work. If these 580,000 people were included in the labor force, February's unemployment rate would have risen to 8.8%.

From June 1976 to July 1976 the unemployment rate increased from 7.5% to 7.8%. Good news? Possibly. From June to July 700,000 people entered the labor force. Of these, 400,000 found jobs. The remaining 300,000 did not find jobs. Thus business was able to employ 400,000 new people. (See Example 11 of Section 6.)

One of the more perplexing aspects of statistical doublespeak is that there is often statistics which support contradictory positions on a subject. For example, consider the contradictory findings of the Gallup and Harris polls in late 1975. Harris found President Ford trailing Senator Hubert Humphrey 52% to 41% when they were matched in a head-to-head race for the presidency. Gallup, however, put President Ford ahead 51% to 39%. This striking contradiction may be explained by the timing of the two surveys. The Harris poll came <u>after</u> Ford announced a Cabinet shake-up and <u>after</u> Ronald Reagan announced he would oppose the President for the 1976 Republican presidential nomination but <u>before</u> Ford traveled to China. The Gallup poll

came after Ford left Peking. Even though foreign travel often results in an increase in popularity for a president, this shift is unusually large. It signals an unstable public opinion of President Ford and a continuation of contradictory findings by national polls through the 1976 presidential campaign.

I have found through my efforts at trying to teach about statistical doublespeak that students find that they can detect this type of doublespeak if they will just try. The major objective of any course or unit on statistical doublespeak is, therefore, to dispel the fear students have of numbers. This is not an easy task but necessary and possible.

There are two keys to getting students to think about statistics:

i) Talk to them in their language - don't force them to learn the nonverbal language of statistics. That is, don't use any mathematical formulas.

ii) Use examples from the media - preferably examples students bring in.

Since this unit is being written for English teachers who will be teaching about doublespeak (verbal and nonverbal), I'm sure you'll be able to refrain from using mathematical formulas - if I can! Taking examples from the media can also be done. The purpose of this unit is to tell you (and, hence, your students) what to look for.

I am presently writing a textbook which would be suitable

for a course on statistical doublespeak but this topic is much too immense for a short unit on nonverbal doublespeak in a course which will necessarily (and rightfully so) put greatest emphasis on verbal doublespeak. For this reason, I will be talking of doublespeak which results from surveys - opinion polls, election polls, TV viewing and radio listening surveys, etc. I've decided to talk about survey doublespeak because, first of all, examples of this type of statistical doublespeak (unfortunately) abound. Secondly, survey doublespeak is probably the type of statistical doublespeak which we can most easily discuss as an "art" without getting involved with "scientific" considerations.

Our approach will be to

i) talk about how surveys can do wrong.

ii) discuss how to evaluate surveys without getting into the strictly "statistical" aspects of survey evaluation

and iii) look at examples so you will be able to aid students in their search for survey doublespeak.

2: TERMINOLOGY.

Let us first take a few minutes to discuss what we mean by statistic, in particular, survey statistics.

There often is interest in the proportion of people in a given group who satisfy some criterion. For example, in pre-election polling we would like to know the proportion of

registered voters who plan on voting for a particular candidate on election day. For a TV viewing survey we wonder what proportion of viewers are watching a particular program. (This proportion is called the audience share of a program.) Other examples include opinion polls. Here we seek the proportion of people who think the President is doing a "good job" or the proportion of people who think this way or that way on most any subject (often to great boredom).

Whatever group of people we are interested in, it is quite often economically and/or practically impossible to talk to everyone in this group to find the proportion we are interested in. (The word "census" is used to describe an attempt at talking to the entire group of interest.) So we talk to only a part (called a sample) of our target group (called a population). Since we are interested in the proportion of people in the population who satisfy some criterion, we will look at the proportion of people in our sample who are of this type. This number is called a statistic.[1]

Statistics we have already alluded to are the proportion of registered voters in a sample who plan to vote for a particular candidate, the proportion of TV viewers in a sample who are watching a particular TV program, etc.

[1]Generally, a statistic is any number obtained from a sample.

A statistic is called inferential when it is used to infer to the corresponding population number. How well the sample represents the population is, then, a key to how "good" the statistic is. So, how do we judge if a sample is representative of a population and, more generally, how "good" a statistic is? Let's see.

3: POOR PRE-ELECTION POLLS.

During an election year many pre-election polls are taken. Some of these polls are taken by the candidates themselves. In this instance we get very little information on the results of the survey and even less of the information we need to judge the credibility of the survey. (Keep this in mind while reading the next section.) At best we get only the results favorable to the candidate. Other pre-election polls are prevalent: Newspapers, newsmagazine, as well as such polling organizations as the Gallup and Harris groups get into the survey business during an election year.

Pre-election polling is unique to inferential statistics. A sample of adults who are registered to vote form the basis for inference to the population of adults who will vote on election day. This type of inference is unique in that the unknown number of interest, the proportion of votes cast for

a candidate on election day, will become known.[2] These polls are then judged by the public which knows exactly how accurate each survey was. There are therefore many examples of bad pre-election polls, some of which you have probably heard about. For example, there was the 1936 election of F.D.R. over Alfred E. Landon and the 1948 election of Harry Truman over Thomas Dewey. In both of these cases pre-election polls did not properly predict the winner. The United States does not, however, have a premium on inaccurate pre-election polling. In Canada, Elliot Trudeau was elected Prime Minister in 1974 contrary to what the polls predicted. Edward Heath's 1970 election over Harold Wilson in Great Britain was not unlike the 1948 U.S. polling disaster. We will consider the polling for the 1936 and 1948 U.S. Presidential Elections in some detail. These examples should aid us in evaluating sample surveys generally.

For twenty years prior to 1936, the Literary Digest was relatively successful in predicting the winner in presidential

[2]Also observe that the population of interest (people voting on election day) does not exist at the time of the survey. A pollster can only ask if a registered voter intends to vote but these people may end up not voting. Besides trying to determine if a person is likely to vote, a pollster must try and determine if a person is likely to change his or her mind before election day. A politician will pay a lot of money to a pollster who can accurately forecast the winner of an election - This is not an easy task.

elections. Up to 20 million cards were mailed with as many as 3 million returns giving the respondents preference in the upcoming election. Based on these large sample sizes of about 3 million the Digest predicted a winner. In a July 12, 1936 newspaper article, George Gallup predicted 56% of the vote for F.D.R. over A.E. Landon indicating that the Digest would be wrong in their upcoming polling attempt! Wrote the Digest:

> "Never before has anyone foretold what our poll was going to show even before it started. Our fine statistical friend (George Gallup) should be advised that the Digest would carry on with those old fashioned methods that have produced correct forecasts exactly one hundred percent of the time."

The 1936 Digest poll was, as Gallup predicted, a disaster. The Digest predicted that Franklin D. Roosevelt would receive 49.9% of the popular vote and 161 out of 531 electoral votes. FDR actually received 60.2% of the popular vote and 523 votes in the Electoral College!

The FDR landslide ended the Digest's pre-election polling and contributed to its demise a few years later. There are two major reasons why this pre-election poll was such a disaster. First, observe that this is a voluntary response survey. People choose to be in the sample by mailing back the postcard sent them by the Digest. In the 1936 poll about 23% (2.3 million of 10 million) of those receiving cards returned a "ballot." Hence, 77% of the "sample" did not respond.

Secondly, the list of names from which the 10 million cards were sent were taken from subscription lists of magazines, telephone directories and automobile owners lists, etc. Such lists were biased toward people of higher incomes. This is especially true since we are talking of depression times.

Since the more educated are more inclined to return questionaires, the sample was also biased toward this group. Hence, even though the sample was of substantial size, nearly 2.3 million for the 1936 poll, it was not representative of the population. This sample more reflected the opinions of a higher income, more highly educated segment of the population and was not representative of the opinions of the general voter population. Gallup saw this bias developing in the Digest's "straw vote."

It should also be noted that this example shows that the sampling procedure is as important if not more important than the size of the sample. This can be seen by comparing the Digest's sample of 2.3 million to present-day national polls of about 1200 people.

Another pre-election polling disaster occurred in 1948. From the time of the last Literary Digest poll of 1936 until the Truman-Dewey Campaign of 1948, the Callup and other polling organizations experienced much success and public support in the pre-election polling arena. But in 1948 the major polling organizations predicted a Thomas Dewey victory and suffered a severe setback in public acceptance when

Harry Truman was elected. We see from Table 3.1 that the major polling organizations errored from 5% to nearly 12% in predicting Truman's vote.

TABLE 3.1 Percentage of Presidential Votes in 1948

	Dewey	Truman	Thurmond	Wallace	Other
National Vote	45.1	49.5	2.4	2.4	.6
Pre-election Polls					
Gallup	49.5	44.5	2.0	4.0	-
Crossley	49.9	44.8	1.6	3.3	.4
Roper	52.2	37.1	5.2	4.3	1.2

The first of the two major reasons for the failure of these polls was, as in the 1936 example, the fact that the sample was not representative but biased toward the educated. Table 3.2 illustrates this point.

TABLE 3.2 Sample Percentages

Education (last school attended)	Population estimates	Gallup (October 14 sample)	Roper (October 25 sample)
Grade School or less	43.5	35.3	27.5
High School	43.4	46.8	48.8
College	13.0	17.9	23.7

The second and probably the most important reason for the poor showing of the pre-election polls of 1948 was the failure of the pre-election polls to detect shifts in voting intentions during the later stages of the campaign. We note (see footnote 2)

that the population of interest in a pre-election poll does not exist at the time of the poll is taken. The population of interest consists of the people who vote on election day. Therefore, the closer to election day that a sample is taken, the more likely the sample is to be representative of this nonexistent population. Roper declared, "As of this September 9, my whole inclination is to predict the election of Thomas E. Dewey by a heavy margin and devote my time and efforts to other things." Gallup and Crossley conducted their last surveys two weeks prior to the election. Yet, 1 in 7 voters decided who to vote for in the last two weeks of the campaign, with 3 out of 4 deciding to vote for Truman. Hence, Truman gained support necessary for victory between the time of the last poll and the actual election.

Since 1948, pre-election polls try and detect voter shifts near election day. You will notice that the last poll before an election is now made public the day before the election with polling conducted a few days before. This last poll is usually a telephone poll that may be completed in one day, the Sunday before election day. It is thought that in this way the population sampled will be more like the population of interest.

Another point of interest, which is closely tied to our discussion of late voter shifts, concerns the 1948 pre-election pollsters' handling of the "UNDECIDED" voter.

There are people who have not decided who they will vote for at the time of a pre-election poll. For the 1948 campaign, the people polled who were undecided were allocated to the candidates on a proportional basis or were eliminated from the sample. This undecided group was substantial prior to the 1948 election - up to 15% (the 1 in 7 previously mentioned). Note today how carefully the "UNDECIDED'S" are watched. (See Example 3 in Section 6.)

In retrospect, the pollsters of the 1948 election had backed themselves into a corner. Their press releases indicated that they were so positive that Dewey would win that they left themselves no out. Assimilators of statistical information should take note: There is uncertainty in inferential statistics.

As an aside, note that not all pollsters were wrong in 1948. Consider the following:

> "One correct poll: Staley Milling Co., Kansas City... is blushing this one off - not because it was wrong, but because it was right.
>
> "Just before the election the company held an informal poll in which farmers cast their votes by calling for a chicken-feed sack with a donkey or an elephant on it. After three weeks, the results stood: 54% for Truman, 46% for Dewey. Unable to believe that it saw, Staley called the thing off. Quipped on executive at the time (with more foresight than he realized), 'If pullets were ballots, President Truman would be a shoo-in for re-election' (Business Week, November 13, 1948, p. 26)."

4: JUDGING A SAMPLE SURVEY.

As mentioned before, a sample is a part of a larger collection or population. The sample is supposed to represent the population. Information from the sample is used to make inference (inferential statistics) to characteristics of the population.

Let us see what we should consider when judging inferential (survey) statistics which we encounter in the media. Included are pre-election polls, public opinion polls, marketing surveys, television viewing surveys, etc. Behind the reported statistics, usually percentages, are many considerations of interpretation which are for the most part subjective, that is, a personal or individual determinations. We do not, therefore, look for the right way to interpret a survey but a good way to interpret a survey.

The source of the survey is, as with the proper understanding of any statistics, important. There are many varied sources of surveys. The best known are the polling organizations of Gallup, Harris and Yankelovich. (See Wheeler, Chapter 3.) The Census Bureau is another important source of survey statistics. A complete list of polling organizations would be long indeed. I would not attempt to try and list "good" or "bad" sources of survey data. We will concentrate on judging the quality of the survey itself. Since such a judgment requires specific information,

the availability of this information serves as a good criterion for judging the source of the survey. A good source will properly describe the survey results with their limitations as well as describing exactly how the data were obtained.

Such a mode for judging the source of a survey is emphasized by the American Association for Public Opinion Research (AAPOR) which in 1968 adopted a code of professional ethics. This code requires that a report of survey results include:

i) the size of the sample and sample design (These are the "scientific" considerations which we will not be investigating.)

and ii) the population, the method of contact for interviewing, time of interviews, and the exact wording of questions.

Not required by the AAPOR is the response rate of a survey. We will discuss this as well as the influences in ii) above in this section.

A caution with regard to the source of a survey before we continue. Never judge survey results solely by the source. As I was investigating a claim by a student newspaper that 9 out of 10 students read the paper in question, I was led to the man in charge of advertising. His sole defense of the results of the survey which was the basis of their advertising claim, even as I probed the limitations of the survey, was that the research company which ran the survey had a "good reputation" in the advertising field. At this point in our discussions, don't judge a survey solely by its source!

The first bit of important information required so we may judge a survey is the population sampled. We will have to determine if the sample is representative of that population as well as being sure that inference is made only to that population. The mechanics of drawing a "credible" sample may be followed so that the sample is representative of some population, however, inference may then be incorrectly made to another (usually larger) population. For example, a credible study may be conducted at some university. The sample of students on which inference is to be based may have been properly drawn and inference may then be properly made to the entire student population at the university under investigation. A serious error would be made if inference is instead made to all college students in that state or even in the country. Inference can be made only to the population actually sampled.

Here lies a source of statistical doublespeak mentioned before: Misusing a valid (survey) statistic. Doublespeak results when survey results are proposed to be representative of a population which is larger than the one actually sampled.

An example of this type of misuse of statistics appeared in a student newspaper. The following comment was made: "...among the 70 percent non-replys (to a questionnaire on abortion) we can be 95 percent certain that less than 10 percent

of Fayette County support no change (in present abortion laws)...." Inference is being made here to individuals not returning a questionnaire. Quite clearly inference cannot be made to individuals not contacted.

Our discussion turns to another backstage influence on survey statistics - nonresponse. A (sometimes large) portion of a sample will not participate in a survey. Reasons for such nonparticipation includes refusal to answer questions or an interviewer's inability to contact a potential respondent. These nonrespondents are "representative" of a part of the population for which there is no information. For example, if 30% of a sample do not respond, we have no information on the 30% of the population which they were to "represent." Inference can be made only to the part of the population represented by the sample (70% of the whole population in this case). Inference to the entire population is not proper. Here lies one of the greatest problems on the proper interpretation of survey results: We are often not told response rates. One needs to know. The AAPOR has not helped in this regard.

Surveys with high nonresponse rates are so-called voluntary response surveys. In such surveys people are not choosen to participate but, rather, people decide whether or not they want to be in the sample. The most common type of voluntary response survey is the mailback survey. Here a questionnaire is mailed to

people or printed in a newspaper. Readers of the questionnaire then choose to be in the sample by mailing in a completed questionnaire. An important aspect of interpretation of such results is, "What 'population' do the respondents represent?" There is usually no answer to this question.

There is only one way to get credible results from a voluntary response survey. Those who do not return the questionnaire need to be contacted (a subsample would suffice) by telephone or through a personal interview. If this is not done, the survey results must be viewed with a suspicion level commensurate with the nonresponse rate. (See Example 9 of Section 6.)

Another important bit of information required to judge a survey is the method of contact used to collect data. For example, respondents may have been contacted by telephone. This necessarily reduces the population or "public" from which opinion has been obtained to people with access to a telephone. Since many telephone surveys select telephone numbers from a directory, the population is further reduced to people with listed telephones. This can be especially restrictive in metropolitan areas where as few as 50% of telephone numbers are listed in a directory. The opinion of people in "unlisted" households can be quite different from the opinion of people in "listed" households. Such differences will vary from question to question but need to be considered.

Be careful in this regard. Participants in a television viewing survey or product research survey may first be contacted by telephone and asked to keep a t.v. log or test a product in their home. Such a technique reduces the population to households with (listed) telephones as in the case of an actual telephone survey.

Another backstage influence on survey results is the timing of a survey. For example, Gallup reported a congressional "approval" rate of 48% in August which is a large increase over the April level of 30%. Such as unusually large increase in the public's view of Congress is put in proper perspective when one realizes that the time was 1974 and between April and August the House Judiciary Committee had voted articles of impeachment against then-President Nixon.

So important was the consideration of the timing of television viewing and radio listening surveys that the Federal Communications Commission proposed rules prohibiting

i) station contests within four weeks of a rating survey,

ii) mass mailings within four weeks of a survey,

and iii) contests where a viewer answers the phone with anything besides "Hello" three weeks before a survey.

Be careful when considering the timing of a survey. Pollsters often make excuses for a poorly designed survey by saying that their results were an accurate measure of opinion for the time the survey was conducted. There may be other reasons for questionable survey results or contradictory findings of different surveys. Consider all the influences mentioned in this section.

We continue our look at surveys by a consideration of another important backstage influence on survey results, the question. Some questions are so constructed as to be more a statement than a question. Interesting "questions" I've run across include: "Should the insurrection of long-haired college students be stopped now or later?" or "Would you favor a no-fault insurance plan which would increase your premiums 30%?"

More subtle might be a question like "Would you favor or oppose deregulation of oil produced in the U.S. if this would encourage development of oil production here at home?" One could well be "opposed" to deregulation of oil but "in favor of" increased oil production at home. How should one interpret an "in favor of" response to the entire question?

Another important aspect of question design is whether the question is open or closed. An open question allows a person to answer in his or her own words. For example, "What do you think of President Ford's performance in office?", is an open question.

A closed question requires a person to pick one of two or three possible answers. For example, "Do you approve or disapprove of President Ford's performance in office?", is a closed question. It is often very hard to answer a closed question. (You should try and answer the closed questions asked in opinion surveys which you read about in the media.) The media, however, prefers to report the (magic) percentage response to closed questions so that this is the type of question which you are most likely to read about.

Many surveys may need no further consideration after consideration has been given these backstage influences: the source, the population, the response rate, the method of contact, timing, and the wording and type of questions asked. For example, a mailback questionnaire encouraged people in a cover letter to respond so that our representatives in Washington will know how much "we" oppose our country becoming "second best to the Soviet Union in military strength." Those choosing to respond were then asked the question, "Should the U.S. have military strength greater than that of the Soviet Union?" No further consideration of the credibility of this "survey" is needed.

5: MORE ON SURVEYS AND SURVEYING.

A Gallup Interview

One can read about national polling organizations as Gallup's American Institute of Public Opinion or even read about polling in general (see the Reference Section) but actually experiencing an interview by a Gallup representative is an interesting experience. I know - it happened to me.

In May of 1975 I was interviewed for a Gallup Survey and would like to share this interesting experience with you. The interview itself lasted about 30 minutes. Questions covered such diverse topics as abortion, luggage, banks and banking, wines, micro-wave ovens, potential candidates for the upcoming presidential election (a year-and-a-half away), advertising in the media, corporate relationships to consumers, lawn mowers, sinus congestion as well as the usual demographic questions on age, income, education, etc! Some of the questions asked would find their way to Gallup's newspaper column. Most of these questions, however, were asked for firms or individuals who contract with Gallup for the service. The term "piggy-backing" is used to describe this over-loading of a questionnaire. Quite clearly one does not have time to do much thinking about the questions asked as you are continually being asked to change thought processes.

More interesting to me and more relevant to our discussions was <u>my interview of the Gallup interviewer</u>.

The young lady worked for an interviewing service located in a nearby metropolitan area. She was paid $2.25 per hour and 12¢ a mile for her work. After the first couple of Gallup surveys she received questionnaires and necessary instructions directly through the mail.

The interviewer had been asking questions for a different Gallup survey each week for the last 10 or 11 months! Each survey was conducted on a separate block in the same general area of our city. That is, each week, usually on Saturday, she would ask a different set of questions (a different Gallup survey) of a member of households on a different block in the same area of town.

Her instructions were to interview 5 adults on a block for each week's survey. One week she'd talk to 3 males and 2 females, the next week 2 males and 3 females. Occasionally she would be instructed to talk to a "special" 6-th person, say, a 16 year-old male.

She was instructed to start at a particular corner of our block, the east corner, go to the 5-th household and then to every household until her quota of 5 interviews was filled. (A household is a "living quarters" for a family or individual(s). The building at the east corner of our block is a four-plex so the 5-th household was the resident in the next house going "clockwise.")

Each household visited would be classified as "NOT HOME," "REFUSED TO BE INTERVIEWED," "COMPLETED INTERVIEW," etc. The interviewer I talked to said that she would have to contact 6 or 7 adults to get the required 5 interviews According to her, some less experienced interviewers may have

a harder time getting 5 completed interviews. (If an interviewer visited all households on a block, he or she would have to revisit households until 5 interviews were completed.)

I was informed that it was not unusual to be "checked." Since one of the questions asked was my phone number, I was told that someone might call to confirm that the interview had indeed taken place.[3]

There were other reasons for getting my phone number as I found out a few months later. A different Gallup representative called to conduct a second interview. This telephone interview was 20 minutes long and involved questions only about wine!

My second interviewer was a male who was making his first series of interviews. He answered an ad in a local newspaper, took a "test" (a trial interview) and then waited until he was contacted to work.

There is a second way to use telephone numbers acquired via a personal interview survey: These numbers form the basis for a subsequent telephone survey. If a person interviewed has a telephone number of, say, 296-5289, a second survey may be conducted by telephone using a number of the form 296-52 -- where the last two digits, not 89 as in the original telephone number, are

[3]One of the difficulties with personal interview surveys is that one cannot be sure that an interview actually took place. An advantage of telephone surveys is that the interviewing can be monitored.

randomly selected. The first five digits are held fixed since, if these numbers are assigned, all such numbers are likely to be assigned. Also, telephone numbers with the same first five digits are likely to be in the same area of a telephone district.

There are two reasons why generating random telephone numbers is not desirable:

i) Time is wasted trying to complete calls using unassigned numbers. (Our telephone company does not always indicate that a number is no longer in service.)

and ii) There is a high nonresponse rate for people with unlisted numbers. These people are so upset that a stranger called their unlisted number that they hang up in a "huff."

In metropolitan areas where as many as 50% of the telephones are unlisted one is forced to use a random number generating procedure even with the disadvantages I've mentioned.

An Arbitron Diary

I have heard many people comment that they have never been asked their opinions for a survey and therefore think that survey results do not reflect their feelings properly. Yet, within a year of being interviewed by the Gallup organization,

I was asked to participate in an Arbition radio listening survey. Arbitron (American Research Bureau) conducts both TV and radio surveys in "local markets." (Emphasis of Neilsen TV Surveys is national.) Local survey markets are generally an enlargement of a Standard Metropolitan Statistic Area (SMSA).[4] The total survey area is the SMSA plus contiguous counties which receive "strong" (radio or TV) signals (at least 2 for radio surveys) from stations in the metro area.

A letter was sent by Arbitron indicating that we would be called and asked to participate in a radio survey. This pre-mailer was to assure us that no one was trying to sell anything. (A pre-mailer of this type is not uncommon as some businesses start a sales pitch with, "I'm conducting a survey." Many people will consequently refuse to talk to a representative of a legitimate survey organization for fear that half-way through the "questions" one would hear something like, "Oh, your car is four years old; would you be interested in a new car?" Licensing of surveys are consequently required in some cities.)

[4]Basically, an SMSA is a county containing a central city of about 75,000 people along with all contiguous counties tied to the central county by certain social/economic criteria. (Eg., 20% of labor force commutes to central city to work.)

A few days after receiving this letter, we were contacted by telephone. After agreeing to participate, a diary was sent each member of our household over 12 years of age. (For a TV survey a diary is sent for each TV set.) Enclosed with a letter of appreciation was a "token" of 50¢ for each participant from our household.

Before the survey week began we were again contacted. This call verified the arrival of the diaries and reminded us to keep an accurate record of our listening habits for the survey week - April 15 to April 21. On the next to the last day of our survey week we were again contacted and reminded to mail in our completed diaries on the 22nd.

I found it quite difficult to remember to record in the diary each and every time I listened to the radio: sometimes at home, other times in the car, or even at a business. (A car dealer must have thought me strange when I asked what radio station they had coming over their speaker system.)

All in all, the diary was probably a good indication of my radio listening habits for that week considering:

i) Easter fell on April 18 so we had out-of-town company most of the week. (Timing of a survey is important.);

and ii) On the first day of the survey week our TV "went on the Fritz" so we did not have a working TV during the entire week. (We therefore represented people with TV sets which were temporarily out-of-order.)

(With regard to our "Easter" survey week: This particular market is surveyed for radio listening once a year. The survey period is 4 weeks in length although an individual keeps a diary for only 1 week. The fact that Easter fell during our week does not mean all diaries traversed this holiday.)

How accurate is an Arbitron survey?[5] Consider:

i) Selection of households is from telephone directories except in areas of high concentrations of blacks or Spanish-speaking peoples. (Complaints that these peoples were not properly represented in telephone directories prompted this exception.)

ii) Diary-type surveys have low response rates. (For a radio survey in our area about 50% of people contacted apparently return useable diaries - diaries which are complete and properly filled out.)

and iii) The sample size is small. An effective sample size of about 600 in an area of about 500,000 people age 12 or more. (How many are listening to a radio at a particular time?)

The written report of the listening habits in an Arbitron market carefully lists the limitations of the survey. The important question remains unanswered: "How representative of the people 12 years or older in a particular market area are the survey results?" That is, how representative of people 12 years and over are the people contacted who agree to participate and return a useable diary?

Also, recall from our discussions in the last section that the F.C.C. has asked that station "contests" not be run during

[5](See Wheeler, Chapter 10 for a discussion of Nielsen TV ratings.)

survey weeks. Arbitron says that if such a contest is run, a notice will be placed on the report of the survey results. I am interested in seeing this report as one station was running a "cash call" contest (guess a 3-digit combination and get up to $1000 in merchandise).

6: EXAMPLES

1.) Let us look at one of the many examples of contradictory findings of surveys which you're likely to find in the media. A Gallup survey reported that the public opposed federal aid to New York City 49% to 42% (late 1975). A telephone survey run jointly by CBS News and The New York Times found the public in favor of aid 55% to 33%. The difference: The Gallup Poll (probably personal interview) was run before a speech in which President Ford spoke out against aid. The CBS-N.Y. Times poll came after this speech. The method of contact as well as the timing of these two polls could explain the doublespeak.

2.) I'm always wondering if our elected officials "survey" public opinion: From a U.S. Senator

"...Each June, I send the questionaire to all the individuals and households on my newsletter mailing list. This year that amounted to 208,300; of that number, just about 22,000 returned the form. As they came into the office the questionnaire were counted, the number logged and the forms kept separate by day. When the number being returned fell to just a few daily, 10% of each day's questionnaires were pulled at random and the responses tallied. From that sample the statistical analysis shown in the enclosed newsletter was made ...".

Here we have a very credible sample of respondents to a questionnaire. This Senator can make inference only to the 22,000 who responded. However, inference cannot be made to all voters in the state or even to the 208,300 who were sent questionnaires. This is because the 208,300 sent a questionnaire are not representative of the voters in the state. People on such lists are party workers and/or people who have written the official. Also, the 22,000 respondents are not representative of the 208,300 sent questionnaires. This is a voluntary response survey. (Note the less than 11% response rate.)

3.) Pre-election polling for the 1968 election is an interesting example to consider. Following the chaotic Democratic Convention in Chicago, the democratic nominee, Hubert Humphrey, trailed the republican candidate, Richard Nixon, by 16%. By late September the margin had changed little. By early October the gap narrowed to 12 points. The third week of October found the difference to be 8%. The last poll, reported the night before the election, labeled the election "too close to call." (Final margin was about 1%.) We see here how important it is to keep a close eye on the "trends" of public opinion. This is what the pollster failed to do in the 1948 Truman/Dewey election.

4.) Recall our discussion of audience share - the percentage of the TV viewers watching (listening to) a TV (radio) program. In reports put out by survey groups the "share" for public TV stations is usually denoted "*" or "-" meaning their share is too small to report. So when our local public TV station got a reported share of 100% for "Man: The Incredible Machine" I was naturally curious. TIMING: That evening all commercial networks in our area were "knocked-out" by an electrical storm. This public TV station was the only station operating. (The public TV transmitor is located outside our county.)

5.) The local telephone company mailed a questionnaire to all of its customers. Many politicians do the same thing - mail questionnaires to all or most of their constituents. This is a public relations gimmic. Such "surveys" have very low response rates (recall example 2). A credible sample survey (telephone or personal interview) would give more accurate information on public opinion at a smaller cost. The questionnaire is sent to give the receiver a feeling of "importance" - not to get a true reading on the public's feelings on different issues.

6.) Consider three surveys which reported:

i) 38% of Americans wanted (ex-President) Richard Nixon to remain in office (Time-Yankelovich telephone survey run May 15th and 16th, 1974),

ii) 41% (Harris poll conducted the week of May 13th to 17th)

and iii) 55% (Questionnaires mailed from the White House were returned between April 29 and May 10. Of the 6000 newspaper editors, broadcasters and White House supporters sent the questionnaire 1,677 responded.)

The Time-Yankelovich and Harris polls are targeted to different populations - adults with access to a telephone and all adults, respectively. One can show however that the differences reported are just a sampling functuation and we have no reason to doubt that the two populations feel the same on this issue.

The White House poll has different results from the other two polls for obvious reasons: Those sent the questionnaire were likely supporters of ex-President Nixon. Also, those returning the questionnaire are likely to be in favor of his remaining in office since the questionnaires were sent from and returned to the White House.

7.) Redbook magazine published a questionnaire asking its (female) readers to respond to certain questions. Over 100,000 responses were received on some very probing questions about sexual beliefs and behavior.

Since only Redbook readers got a questionnaire and they

volunteered to be in the sample by mailing in answers to these questions, this survey is not too dissimilar from the 1936 Literary Digest pre-election poll of 1936. (Recall our discussion in Section 3.)

Be careful in instances like this. Redbook claimed their sample was representative since the percentage of respondents in different demographic categores (age, marital status, religion, etc.) were close to population percentages. Warning: People who choose to respond (especially to this October 1975 sex survey) may be a very select part of the (female) population of interest. Respondents in this example may represent, for example, females who are open about their sexual behavior. (Also, how truthful are respondents likely to be in such a situation?)

8.) A (hearsay) response by a U.S. Senator's aide indicated that this official kept his hand "on the pulse" of our area by reading a "Your Opinion Counts" column in a local newspaper. Each Saturday unedited questions which were sent in by readers were published. ("Do you think Congress is going its usual crummie job?" - Spelling was not even corrected.) The following week percentage responses of those who mailed in answers were reported. (Percentages were reported but not the number responding. Hence, 20% could be 1 in 5 or 1000 in 5000.)

I don't think that I need to dwell on this one. It is quite sad, however, that a U.S. Senator (apparently) saw value in such a straw vote.

9.) The response rate for a survey is an important bit of information which is often hard to obtain as we have discussed. People are so harassed on the telephone that a credible polling group has little hope of getting a high response rate.[6] As little as 50% response can be expected for telephone surveys in metropolitan areas.

For mailback surveys, pollsters try to increase response rates by follow-up mailings to those who haven't yet responded. This means that the pollster must know who has responded making confidentiality of responses a problem. A national newspaper had a survey organization conduct a mailback survey for them. This organization used invisible ink to try and secretly identify respondents!

10. As a class project have students ask both open and closed questions on some thought - provoking topic. How hard is it to get closed responses? On the other hand, how difficult would it be for a newspaper to report responses to an open question?

11. Why do you think 700,000 people entered the labor force between June 1976 and July 1976? (Recall our discussion of this in

[6] I've been called for 3 surveys in the last two months. This doesn't include the surveys I discussed in Section 5.

Section 1.) There are two possible explanations:

i.) Since the economy was improving, people felt that it was worth their effort to seek employment. (Recall that a person enters the labor force by trying to find a job.) and/or ii.) Inflation made it so hard for families to live on one salary that wives started to look for work.

Both explanations are likely to be relevant. The news reports noted the large proportion of housewives entering the labor force at that time. The very fact that 400,000 people found jobs indicates an improved labor market. (Improvement over the February 1975 figures discussed in Section 1 is clear.)

7: REFERENCES

As teachers we always like to be one step ahead of our students (as hard as this sometimes is). There are a few good - and readable - books which may help in this regard as you teach about statistical doublespeak.

With emphasis on surveys there is:

i) Gallup, G. (1972). The Sophisticated Poll-Watchers Guide. Princeton Opinion Press.

ii) Roll, C. W., Jr. and Cantril, A. H. (1972). Polls: Their Use and Misuse in Politics. Basic Books.

and iii) Wheeler. M. (1976). Lies, Damn Lies and Statistics. Liveright.

On the more general topic of statistics and statistical doublespeak consider:

i) Bross, I.D.J. (1957). Scientific Strategies in Human Affairs: To Tell the Truth. Exposition Press.

ii) Campbell, S. (1974). Flaws and Fallacies in Statistical Thinking. Prentice Hall.

iii) Haack, D.G. (January 1978). A stricklyyverbal statistics text. Duxbury Press.

iv) Hauser, P.M. (1975). Social Statistics in Use. Russell-Sage Foundation.

v) Huff, D. (1954). How to Lie with Statistics. Norton.

vi) Messick, B.M. Mathematical Thinking in Behavioral Sciences. Readings from Scientific American. Freeman.

vii) Reichard, R. (1974). The Figure Finaglers. McGraw-Hill.

and viii) *Taner, J. (editor) (1972). Statistics: A Guide to the Unknown. Holden-Day.

*An excellent collection of readable essays on the application of statistics in many different areas.